```
DIALOG(R) File 351: Derwent WPI
 (c) 2001 Derwent Info Ltd All rts. reserv.
009240301
WPI Acc No: 1992-367719/199245
XRAM Acc No: C92-163297
Strong transcriptional promoter from Kluyveromyces lactis - provides
 efficient expression of heterologous proteins e.g. human serum albumin in
 yeast
Patent Assignee: RHONE POULENC RORER SA (RHON ); RHONE-POULENC RORER SA
   (RHON )
 Inventor: FLEER R; FOURNIER A; MAYAUX J; YEH P
Number of Countries: 026 Number of Patents: 016
Patent Family:
Patent No
               Kind
                      Date
                               Applicat No
                                              Kind
                                                      Date
                                                               Week
EP 511912
                    19921104
                Al
                               EP 92401206
                                                    19920428
                                               Α
                                                              199245
WO 9219751
                A1
                    19921112
                               WO 92FR375
                                               Α
                                                    19920428
                                                              199248
 FR 2676070
                A1
                    19921106
                               FR 915294
                                               Α
                                                    19910430
                                                              199301
 ZA 9203083
                Α
                    19930127
                               ZA 923083
                                               A
                                                    19920428
                                                              199310
AU 9217892
                    19921221
                Α
                               AU 9217892
                                               Α
                                                    19920428
                                                              199311
                               WO 92FR375
                                               Α
                                                    19920428
NZ 242543
                Α
                    19931026
                               NZ 242543
                                               Α
                                                    19920429
                                                              199345
FI 9304806
                Α
                    19931029
                               WO 92FR375
                                               Α
                                                    19920428
                                                              199402
                               FI 934806
                                               Α
                                                    19931029
NO 9303654
                Α
                    19931011
                               WO 92FR375
                                               Α
                                                    19920428
                                                              199404
                               NO 933654
                                               Α
                                                    19931011
 EP 584166
                A1
                    19940302
                               EP 92910339
                                               Α
                                                    19920428
                                                              199409
                               WO 92FR375
                                               Α
                                                    19920428
 JP 6506602
                    19940728
                               JP 92510457
                                               A
                                                    19920428
                                                              199434
                               WO 92FR375
                                                    19920428
                                               Α
HU 67448
                Т
                    19950428
                               WO 92FR375
                                               Α
                                                    19920428
                                                              199523
                               HU 933087
                                               Α
                                                    19920428
AU 658630
                В
                    19950427
                               AU 9217892
                                                Α
                                                    19920428
                                                              199525
 US 5646012
                Α
                     19970708
                               WO 92FR375
                                                Α
                                                    19920428
                                                              199733
                               US 93140093
                                                    19931101
                                                Α
                               US 95483639
                                                    19950607
                                                Α
 EP 584166
                В1
                    19980325
                               EP 92910339
                                                Α
                                                    19920428
                                                              199816
                               WO 92FR375
                                                A
                                                    19920428
 DE 69224902
                Ε
                     19980430
                               DE 624902
                                                Α
                                                    19920428
                                                               199823
                               EP 92910339
                                                A
                                                    19920428
                               WO 92FR375
                                                Α
                                                    19920428
 ES 2113948
                Т3
                    19980516
                               EP 92910339
                                                Α
                                                    19920428
                                                              199826
 Priority Applications (No Type Date): FR 915294 A 19910430
 Cited Patents: 3.Jnl.Ref; EP 361991
 Patent Details:
 Patent No Kind Lan Pg
                           Main IPC
                                       Filing Notes
               A1 F 22 C12N-015/81
 EP 511912
    Designated States (Regional): PT
               A1 F 27 C12N-015/81
 WO 9219751
    Designated States (National): AU CA FI HU JP KR NO US
   Designated States (Regional): AT BE CH DE DK ES FR GB GR IT LU MC NL SE
               Αl
                      31 C12N-015/29
 FR 2676070
 ZA 9203083
               Α
                      34 C12N-000/00
 AU 9217892
               Α
                         C12N-015/81
                                       Based on patent WO 9219751
 EP 584166
               A1 F
                         C12N-015/81
                                       Based on patent WO 9219751
```

4/5/3

```
Designated States (Regional): AT BE CH DE DK ES FR GB GR IT LI LU NL SE
                                 Based on patent WO 9219751
                    C12N-015/67
            W
JP 6506602
                                   Based on patent WO 9219751
HU 67448
             Т
                     C12N-015/81
            В
                     C12N-015/11
                                   Previous Publ. patent AU 9217892
AU 658630
                                   Based on patent WO 9219751
                  23 C12P-021/02
                                   Cont of application WO 92FR375
US 5646012 A
                                   Cont of application US 93140093
EP 584166 B1 F 22 C12N-015/81
                                   Based on patent WO 9219751
  Designated States (Regional): AT BE CH DE DK ES FR GB GR IT LI LU NL PT
DE 69224902
            Ε
                      C12N-015/81
                                   Based on patent EP 584166
                                   Based on patent WO 9219751
           Т3
                     C12N-015/81
                                   Based on patent EP 584166
ES 2113948
                     C12N-015/12
NZ 242543
            Α
FI 9304806
            Α
                      C12N-000/00
NO 9303654
            Α
                      C12N-000/00
```

Abstract (Basic): EP 511912 A

The second section of

New DNA sequences (A) contain all or part of a specified 2249 bp sequence reproduced in the specification or of its complementary strand, or of a deriv. of these. (A) has the activity of a transcriptional promoter. Also new are (1) recombinant DNA (A') contg. (A) and (2) recombinant cells contg. (A) or (A').

Pref., (A') contains 1 or more structural genes plus segments which cause secretion of expression products. Partic. (A') is part of an expression plasmid of the autonomous or integrating types.

USE/ADVANTAGE - (A) is a very efficient promoter in yeasts, esp. Kluyveromyces and is used to express recombinant genes partic. simultaneous expression in 2 opposite orientations. These genes are esp. those of pharmaceutical or nutritional interest, esp. that for human serum albumin (HSA)(

Dwg.0/14

Title Terms: STRONG; TRANSCRIBING; PROMOTE; KLUYVEROMYCES; LACTIS; EFFICIENCY; EXPRESS; HETEROLOGOUS; PROTEIN; HUMAN; SERUM; ALBUMIN; YEAST Derwent Class: B04; D16

International Patent Class (Main): C12N-000/00; C12N-015/11; C12N-015/12;
C12N-015/29; C12N-015/67; C12N-015/81; C12P-021/02

International Patent Class (Additional): C07H-021/04; C07K-019/765;
 C12N-001/16; C12N-001/19; C12N-015/14; C12N-015/66

File Segment: CPI

11) Numéro de publication : 0 511 912 A1

12

DEMANDE DE BREVET EUROPEEN

(21) Numéro de dépôt : 92401206.5

(22) Date de dépôt : 28.04.92

(5) Int. Cl.⁵: **C12N 15/81,** C12N 1/19, C12P 21/02

(30) Priorité: 30.04.91 FR 9105294

(3) Date de publication de la demande : 04.11.92 Bulletin 92/45

Etats contractants désignés :

PT

PT

Représe

Représe

PT

Représe

R

① Demandeur: RHONE-POULENC RORER SA 20, avenue Raymond Aron F-92160 Antony (FR) 1 Inventeur: Fleer, Reinhard
1 Allée Port Royal, Résidence de l'Abbaye
F-91190 Gif Sur Yvette (FR)
Inventeur: Fournier, Alain
28 Avenue Roger Salengro
F-92000 Chatenay Malabry (FR)
Inventeur: Mayaux, Jean-François
21 ter Boulevard de la République
F-92260 Fontenay aux Roses (FR)

Inventeur : Yeh, Patrice 11bis, rue Lacépède F-75005 Paris (FR)

(74) Mandataire: Savina, Jacques et al Rhône-Poulenc Rorer S.A. Direction Brevets (t144) 20 avenue Raymond Aron F-92165 Antony Cédex (FR)

(54) Promoteur de levure et son utilisation.

Cinvention concerne des séquences d'ADN comprenant tout ou partie du promoteur du gène PGK de K.lactis, ou d'un dérivé de celui-ci, et possèdant une activité de promoteur transcriptionnel. Elle concerne également l'utilisation de ces séquences pour l'expression de gènes recombinés.

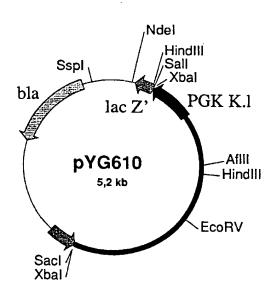


FIGURE 3

10

15

30

La présente invention concerne le domaine de la biologie moléculaire. Plus particulièrement, elle concerne une nouvelle séquence d'ADN présentant une activité de promoteur transcriptionnel, des vecteurs d'expression contenant cette séquence, et son utilisation pour la production de protéines, et par exemple de protéines hétérologues. L'invention concerne aussi les cellules recombinées contenant cette séquence d'ADN.

Les progrès accomplis dans le domaine de la biologie moléculaire ont permis de modifier des microorganismes pour leur faire produire des protéines hétérologues. En particulier, de nombreuses études génétiques ont porté sur la bactérie E. coli. Toutefois, l'application industrielle de ces nouveaux modes de production est encore limitée, en particulier par les problèmes d'efficacité d'expression des gènes dans ces microorganismes recombinés. Aussi, dans le but d'augmenter les performances de ces systèmes de production, des recherches ont été effectuées afin d'isoler des promoteurs forts, permettant d'obtenir des niveaux élevés d'expression de protéines hétérologues. Chez E.coli, on peut citer en particulier les promoteurs des opérons tryptophane et lactose.

Plus récemment, chez la levure S. cerevisiae, des études ont porté sur des promoteurs dérivés de gènes impliqués dans la glycolyse. On peut citer notamment les travaux sur le promoteur du gène de la 3-phosphoglycerate kinase PGK (Dobson et al., Nucleic Acid Res. 10, 1982, 2625; Hitzeman et al., Nucleic Acid Research 1982, 7791), sur celui du gène de la glyceraldéhyde-3-phosphate déshydrogénase GAPDH (Holland et al., J.Biol.Chem. 254, 1979, 9839; Musti et al., Gene 25, 1983, 133), sur celui du gène de l'alcool déshydrogénase 1 ADH1 (Bennentzen et al., J.Biol.Chem. 257, 1982, 3018; Denis et al., J.Biol.Chem. 25, 1983, 1165), ou sur celui du gène de l'enolase 1 ENO1 (Uemura et al., Gene 45, 1986, 65).

Récemment, des outils génétiques ont été développés afin de se servir de la levure Kluyveromyces comme cellule hôte pour la production de protéines recombinantes. La mise en évidence d'un plasmide de type 2-micron originaire de K-drosophilarum (plasmide pKD1 - EP 241 435) a permis d'établir un système hôte/vecteur très efficace pour la production de protéines recombinantes (EP 361 991). Cependant, les promoteurs utilisés dans ce système n'ont jamais été optimisés. En particulier, il s'agit essentiellement de promoteurs hétérologues, c'est-à-dire provenant d'autres microorganismes, tel que notamment S.cerevisiae. Cette situation peut engendrer différents inconvénients, et notamment limiter l'activité du promoteur à cause de l'absence de certains éléments de la machinerie transcriptionnelle (par exemple de transactivateurs), présenter une certainee toxicité pour la cellule hôte due à une absence de régulation, ou affecter la stabilité du vecteur.

Dans ces conditions, le manque de promoteurs homologues forts chez Kluyveromyces constitue un facteur limitant dans l'exploitation industrielle de ce système d'expression.

La Demanderesse a maintenant identifié, cloné et séquencé une région du génome de Kluyveromyces lactis présentant une activité de promoteur transcriptionnel (voir figure 1). Plus précisément, cette région correspond au promoteur du gène PGK de K.lactis. Cette région, ou des dérivés ou fragments de celle-ci, peut être utilisée de manière très performante pour la production de protéines recombinantes chez les levures du genre Kluyveromyces. Il est entendu que cette séquence peut également être utilisée dans d'autres organismes hôtes.

Par ailleurs, l'analyse de la région du génome de Kluyveromyces obtenue a permis de mettre en évidence 2 phases de lecture dans les 2 orientations opposées (voir figure 2). Cette observation indique que le brin complémentaire de la région présentée sur la figure 1 possède également une activité promotrice agissant dans l'autre orientation.

Un objet de la présente invention réside donc dans une séquence d'ADN comprenant tout ou partie de la séquence présentée à la figure 1 ou de son brin complémentaire, ou d'un dérivé de celles-ci, et possédant une activité de promoteur.

Au sens de la présente invention, on entend par dérivé, toute séquence obtenue à partir de la séquence donnée dans la figure 1 par modifications structurales (mutations, délétions, substitutions, additions, fragmentations...) conservant une activité de promoteur. En particulier, les mutations peuvent porter sur un ou plusieurs nucléotides, et les additions et/ou substitutions peuvent porter sur des éléments de régulation, ou des régions activatrices telles que les "UAS".

Lorsqu'un dérivé est réalisé, son activité de promoteur transcriptionnel peut être mise en évidence de plusieurs façons, et en particulier en plaçant sous le contrôle de la séquence étudiée, un gène de résistance ou un marqueur de complémentation. Toute autre technique connue de l'homme de l'art peut bien évidemment être utilisée à cet effet.

Un objet plus particulier de l'invention concerne une séquence d'ADN correspondant à la région comprise entre les 2 phases ouvertes ORF PGK et ORF X, telle que présentée sur la figure 6.

Un autre objet de l'invention conceme un ADN recombinant comprenant une séquence d'ADN telle que définie ci-dessus.

Cet ADN recombinant peut contenir par exemple la séquence promotrice présentée à la figure 1 ou un dérivé de celle-ci, dans laquelle est inséré un site de restriction, facilitant l'utilisation de cette séquence comme promoteur, "portable".

Préférentiellement, cet ADN recombinant contient en outre un ou plusieurs gènes de structure.

15

30

40

45

50

Encore plus préférentiellement, l'ADN recombinant contient également des signaux permettant la sécretion du produit d'expression du ou desdits gènes de structure.

Dans un mode de réalisation particulier de l'invention, l'ADN recombinant fait partie d'un plasmide d'expression, qui peut être à réplication autonome ou intégratif.

En particulier, des vecteurs à réplication autonome peuvent être obtenus en utilisant des séquences à réplication autonomes (ARS) chez l'hôte choisi. Notamment, chez la levure, il peut s'agir d'origines de réplication dérivées de plasmides connus (pKD1, 2μ, etc).

Les vecteurs intégratifs peuvent être obtenus notamment en utilisant des séquences homologues à certaines régions du génome de l'hôte, permettant, par recombinaison homologue, l'intégration du vecteur.

La séquence présentée sur la figure 1 a été obtenue par criblage d'une banque d'ADN génomique total de Kluyveromyces lactis au moyen d'une sonde hétérologue provenant du gène PGK de S. cerevisiae. La Demanderesse a en effet montré qu'il est possible de cloner une région promotrice chez Kluyveromyces, par hybridation à partir de sondes hétérologues correspondant à un gène de S. cerevisiae. Les détails du clonage de la séquence sont donnés dans les exemples. La région intergénique peut ensuite être isolée à partir de cette séquence, notamment par insertion de sites de restriction en utilisant la technique d'amplification par PCR comme indiqué dans les exemples.

Un autre objet de l'invention concerne les cellules recombinées contenant une séquence d'ADN telle que définie ci-avant.

Avantageusement, les cellules sont choisies parmi les levures, et encore plus préférentiellement, parmi les levures du genre Kluyveromyces. Il est entendu cependant que l'invention couvre toute les cellules recombinées dans lesquelles les régions promotrices de l'invention sont actives.

Ces cellules peuvent être obtenues par toute méthode permettant d'introduire un ADN étranger dans une cellule. Il peut s'agir notamment de transformation, électroporation, ou toute autre technique connue de l'homme de l'art.

Un autre objet de l'invention concerne l'utilisation d'une séquence telle que précédemment définie pour l'expression de gènes recombinés.

Comme l'illustrent les exemples, les séquences d'ADN selon l'invention permettent en effet une production à des niveaux élevés de protéines recombinantes.

Par ailleurs, l'activité promotrice bidirectionnelle des séquences de l'invention permet une utilisation particulièrement avantageuse. Notamment, il est possible d'utiliser ces séquences dans les 2 orientations possibles, pour l'expression simultanée de plusieurs gènes de structure.

Avantageusement, l'invention concerne l'utilisation d'une séquence telle que précédemment définie pour l'expression simultanée, dans les 2 orientations opposées, de gènes recombinés.

Avantageusement, les séquences de l'invention peuvent être utilisées pour l'expression de gènes codant pour des protéines d'intérêt pharmaceutique ou agroalimentaire. A titre d'exemple, on peut citer les enzymes (tels que notamment la superoxide dismutase, la catalase, les amylases, les lipases, les amidases, la chymosine etc.), les dérivés sanguins (tels que la sérum-albumine, l'alpha- ou la béta-globine, le facteur VIII, le facteur IX, le facteur van Willebrand, la fibronectine, l'alpha-1 antitrypsine etc.), l'insuline et ses variants, les lymphokines (telles que les interleukines, les interférons, les facteurs de stimulation des colonies (G-CSF, GM-CSF, M-CSF...), le TNF, le TRF etc.), les facteurs de croissance (tels que l'hormone de croissance, l'érythropoiétine, le FGF, l'EGF, le PDGF, le TGF etc.), les apolipoprotéines, des polypeptides antigéniques pour la réalisation de vaccins (hépatite, cytomégalovirus, Eppstein-Barr, herpes etc.), ou encore des fusions de polypeptides telles que notamment des fusions comportant une partie active fusionnée à une partie stabilisatrice (par exemple des fusions entre l'albumine ou des fragments d'albumine et le récepteur ou une partie d'un récepteur de virus (CD4, etc)).

L'invention sera plus complètement décrite à l'aide des exemples qui suivent, qui doivent être considérés comme illustratifs et non limitatifs.

LEGENDE DES FIGURES

Figure 1 : Séquence nucléotidique de la région de 2,2 kb du fragment chromosomique situé en amont du codon d'initiation de la traduction du gène PGK de K.lactis possédant l'activité promotrice.

Figure 2: Analyse des phases ouvertes de lecture. Les demi-traits verticaux représentent des codons d'initiation de la traduction. Les traits verticaux entiers représentent des codons stop. Les régions claires mettent en évidence les phase ouvertes de lecture (ORF X et ORF PGK).

Figure 3: Carte de restriction du plasmide pYG610. La région noire correspond à la région isolée du génome de K.lactis.

<u>Figure 4</u>: Stratégie de séquençage du fragment Xbal de 2,5 kb.

Figure 5 : Séquence et localisation des oligodéoxynucléotides utilisés dans la réaction de PCR, pour l'insertion d'un site HindIII en -6 de l'ATG de la séquence de la figure 1. Les oligodéoxynucléotides sont représentés en italique. L'ATG correspond au codon d'initiation de la tra-

duction du gène PGK.

Figure 6: Séquence nucléotidique de la région intergénique du fragment 2,2 kb. 6(a): oligodéoxynucléotides utilisés dans la réaction de PCR. 6(b): Fragment Sal(I)-HindIII correspondant aux nucléotides 1343 à 2246 sur la séquence de la figure 1.

Figure 7: Stratégie de construction du plasmide pYG45.

<u>Figure 8</u>: Stratégie de construction de cassettes d'expression de la sérum-albumine humaine.

Figure 9: Stratégie de construction du plasmide pYG65.

<u>Figure 10</u>: Stratégie de construction du plasmide pYG70.

Figure 11: Stratégie de construction du plasmide pYG72.

Figure 12: Stratégie de construction du vecteur pYG621.

Figure 13: Mise en évidence par "Northem blot" de l'expression du gène de l'albumine humaine sous la dépendance du promoteur PGK de K.lactis. Les échantillons correspondent à 10µg d'ARN total. 18S et 28S sont les positions des ARN ribosomiques 18S et 28S. ALB = fragments reconnus par la sonde correspondant au gène de l'albumine; URA = fragments reconnus par la sonde correspondant au gène URA A de K.lactis servant de témoin de dépôt.

Figure 14: Mise en évidence de la production d'albumine dans les souches transformées par le vecteur d'expression pYG621 contenant le promoteur PGK de K-lactis. Les échantillons correspondent à 30 μl de sumageant de culture ; les bandes au niveau du marqueur 66 kd correspondent à l'albumine. M = marqueurs de masse moléculaire : anhydrase carbonique bovine (31Kd), ovalbumine (45Kd), BSA (66Kd), phosphorylase b de lapin (92kd).

EXEMPLES

1/ Isolement de la région promotrice du gene PGK de K lactis.

La séquence présentée sur la figure 1 a été obtenue par criblage d'une banque d'ADN génomique total de <u>Kluyveromyces lactis</u> CBS2359 au moyen d'une sonde hétérologue provenant du gène PGK de <u>S. cerevisiae</u> (Dobson et al., Nucleic Acid Res. 1982, 10, 2625). Plus précisément, la sonde utilisée correspond au fragment N-terminal <u>Pvul-EcoRl</u> de 1,3 kb du gène PGK de S. cerevisiae.

En "Southern blot" (Southern et al., J.Biol.Chem., 1975, 98, 503), la sonde utilisée s'hybride avec deux fragments différents lorsque l'ADN génomique est digéré par XBal. L'un d'eux, de 2,5 kb environ, a été isolé par criblage d'une banque génomique restreinte de

K. lactis CBS2359, constituée de fragments d'ADN coupés par Xbal, d'une taille comprise entre 2 et 3 kb, introduits dans le plasmide pUC18 au site Xbal. Une banque de 500 clones a ainsi été constituée, puis criblée avec la sonde hétérologue décrite ci-dessus.

Par hybridation sur colonies, un clone a pu être identifié et son ADN plasmidique a été préparé. Ce plasmide (pYG610) contient un fragment d'ADN génomique de 2,5 kb, dont la carte de restriction est présentée à la figure 3. Le plasmide pYG611 contient le même insert dans la direction opposée (voir figure 8).

Dans une seconde étape, le fragment de 2,5 kb ainsi isolé a été séquencé, en utilisant la méthode de Sanger (Sanger et al., Proc.Nat.Acad.Sci 74, 1977, 5463). Pour cela, le fragment issu de pYG611 a d'abord été sous-cloné dans les bactériophages M13tg130 et M13_tg131. La stratégie de séquençage du fragment est schématisée sur la figure 4.

L'analyse de la séquence obtenue montre que le fragment isolé contient une partie codant pour la région N-terminale de la protéine Pgk de Kluyveromyces lactis (0,3 kb), et 2,2 kb correspondant à la région promotrice située en amont du site d'initiation de la traduction. Elle montre de plus que, dans l'orientation opposée par rapport au gène PGK, se situe une seconde phase de lecture située à environ 0,9 kb en amont de l'ATG du gène PGK (figure 2).

La comparaison de cette séquence avec celle du promoteur du gène PGK de S. cerevisiae fait apparaitre l'absence d'homologie particulière, notamment avec son élément de régulation. Cette séquence correspond donc à une région promotrice entièrement originale, très distincte de celles déjà décrites, sur le plan de sa structure, et par conséquent sur le plan de sa régulation.

2/ Construction de vecteurs d'expression pour la production de proteines hétérologues :

Cet exemple illustre l'utilisation des capacités promotrices de la séquence de 2,2 kb de la séquence de la figure 1 et de séquences dérivées.

a) Insertion d'un site de restriction en -6 de l'ATG.

L'insertion de ce site permet ensuite d'introduire en aval du promoteur obtenu tout gène que l'on désire exprimer. Pour des raisons de compatibilité avec des vecteurs d'expression existant (EP 361 991), des promoteurs "portables" ont été construits sous forme de fragments Sall-HindIII.

Un site HindIII a été introduit en position -6 par rapport au site d'initiation de la traduction (ATG) du gène PGK en utilisant la technique d'amplification par PCR (Mullis et al., Meth.Enzymol. 155, 1987, 335). Dans ce but, 2 oligodéoxynucléotides ont été utilisés, qui sont présentés sur la figure 5.

L'oligodéoxynucléotide A correspond à la sé-

35

45

quence située à 467 pb en amont du codon ATG, au niveau d'un site HindIII, qui sera ainsi remplacé par un site Sall lors de l'amplification. L'oligodéoxynucléotide B correspond à la séquence en amont du site d'initiation, et permet d'introduire un site HindIII en position -6.

Le fragment obtenu par PCR a été inséré entre les sites Sall et HindIII du bactériophage M13tg130 afin de vérifier par séquençage que des mutations ne sont pas apparues lors de l'amplification.

b) Construction de cassettes d'expression de la sérum-albumine humaine : figure 8.

L'ADN recombinant de 474 pb obtenu ci-dessus a été introduit au niveau des sites Sall et Hindlll, dans le plasmide pYG45 (figure 7) pour obtenir le vecteur pYG614 (figure 8). Le plasmide pYG45 contient une cassette d'expression constituée du promoteur et du terminateur du gène PGK de S. cerevisiae entre lesquels, au niveau d'un site HindIII, est inséré le gène codant pour la prépro-sérum-albumine humaine (séquence prépro-HSA). pYG45 est dérivé de pYG18 (voir brevet EP 361 991) par sous-clonage du fragment Sall-BamHI contenant la cassette d'expression HSA, dans les sites correspondants du vecteur plC-20RDH (figure 7), pIC-20RDH est obtenu par digestion du plasmide pIC-20R (March et al., Gene 32, 1984, 481) avec l'enzyme HindIII, remplissage des extrémités avec le fragment Klenow de la polymérase I d'E.coli et recircularisation avec la T4 DNA ligase.

A partir du plasmide pYG614, le fragment Sall-Sacl peut être isolé par digestion. Il contient : une région promotrice dérivée de la séquence de la figure 1, le gène de l'albumine et le terminateur du gène PGK de S. cerevisiae. Il constitue une cassette d'expression qui peut être insérée dans un plasmide pour constituer un vecteur d'expression.

Une autre cassette d'expression peut être obtenue à partir du plasmide pYG614, par clonage du fragment AfIII-Sacl contenant une partie du promoteur PGK de l'invention, le gène de l'albumine (prépro-HSA) et le terminateur PGK de S. cerevisiae dans le plasmide pYG611 décrit préalablement. Ceci génère le plasmide pYG615. Le fragment Sall-Sacl contenant : la région promotrice de la figure 1 entière, le gène codant pour la prépro-sérum-albumine, et le terminateur du gène PGK de S. cerevisiae, peut ensuite être isolé par digestion. Ce fragment constitue une seconde cassette d'expression de l'albumine utilisant la séquence promotrice de l'invention.

c) Construction de vecteurs d'expression de l'albumine.

Des vecteurs d'expression de l'albumine peuvent être construits par insertion des cassettes d'expression obtenues ci-dessus dans des plasmides navettes K.lactis/E.coli tels que pYG72 (figure 10). En particulier, un vecteur d'expression a été obtenu (vecteur pYG621) par insertion du fragment Sall-Sacl de pYG615 contenant la cassette d'expression de l'albumine dans le vecteur pYG72 (voir figure 10). Ce vecteur correspond au plasmide pKan 707 (voir EP 361 991) dans lequel le fragment Sacl contenant le gène URA3 a été éliminé, ainsi que le site unique HindIII présent dans le gène aph pour faciliter les constructions ultérieures. Le gène aph code pour l'aminoglycoside 3'-phosphotransférase (I) (Oka et al., J. Mol-Biol. 147, 1981, 217) et est utilisé comme marqueur de résistance au G418 chez la levure. Le fragment Pstl du plasmide pKan707 contenant le gène aph a été sous-cloné dans le bactériophage M13mp7 pour donner le vecteur pYG64 (figure 9). Le site HindIII présent dans ce gène a été détruit par mutagénèse dirigée selon la méthode décrite par Taylor et al. (Nucleic Acid Res. 13, 1985, 8749). Le plasmide résultant a été appelé pYG65 (figure 9). L'oligodéoxynucléotide utilisé pour la mutagénèse avait la séquence 5'-GAA ATG CAT AAG CTC TTG CCA TTC TCA CCG -3' et transformait le triplet CTT codant pour l'acide aminé 185 (Leu) en CTC. Ce changement ne modifie pas la séquence protéique résultante. Pour construire le plasmide pYG72, la partie contenant le réplican bactérien du vecteur pKan707 a été isolée par digestion avec l'enzyme EcoRI et recircularisation avec la T4 DNA ligase pour obtenir pYG69. Le fragment Pstl présent dans ce dernier vecteur contenant le gène aph a été remplacé par le fragment équivalent muté provenant de pYG65. Cette construction a été appelée pYG70. La séquence de pKD1 de 4,7 kb comprise entre les sites EcoRI et SacI a été introduite dans ce dernier vecteur pour obtenir pYG72. Le vecteur pYG621 (figure 11) a été obtenu par insertion du fragment Sall-Sacl contenant la cassette d'expression de l'albumine provenant de pYG615.

3/ Construction d'une cassette permettant d'utiliser la région promotrice dans les 2 orientations.

Cette construction a été obtenue par introduction d'un site <u>Sall</u> et d'un site <u>HindIII</u> de part et d'autre de la région comprise entre les 2 phases ouvertes de lecture identifiées sur la figure 2: ORF PGK et ORF X, soit au niveau des nucléotides 1343 et 2246 sur la figure 1.

Cette construction a été réalisée par la technique de PCR en utilisant d'une part l'oligodéoxynucléotide A qui introduit un site Sall en position -1 par rapport au site d'initiation de la traduction du gène PGK, et d'autre part l'oligodéoxynucléotide B qui introduit un site HindIII en position -1 par rapport au site d'initiation de la traduction du gène X (voir figure 6(a)). Ensuite, pour éliminer un site HindIII présent dans la région promotrice, 3 réactions de PCR ont été effectuées en utilisant à chaque étape le plasmide pYG610

35

40

45

50

55

comme matrice:

- les 2 premières, pour amplifier les régions de part et d'autre du site <u>HindIII</u> en utilisant les oligodéoxynueléotides A et B couplés respectivement aux oligodéoxynucléotides C et D (figure 6). Ces 2 derniers sont complémentaires et permettent d'introduire une mutation ponctuelle au niveau du site HindIII interne;

- la dernière, en utilisant les 2 produits d'amplification précédents comme amorce, pour générer le fragment final contenant la région promotrice modifiée.

Cette région peut ensuite être introduite dans les vecteurs décrits dans l'exemple 2, et utilisée comme promoteur bidirectionnel.

4/ Expression d'albumine

Le vecteur pYG621 a été introduit par transformation dans la souche K.lactis MW98-8C (CBS 579.88), en utilisant la technique éthylène glycol/diméthylsulfoxyde (Durrens et al., 1990, Curr.Genet. 18, 7). Cette souche dérive de la souche sauvage CBS2359 et présente le génotype : Mat_, uraA, LysA, argA, K+, cir. Les levures transformées sont sélectionnées pour le phénotype "G418-resistant" que confère le plasmide pYG621 sur milieu YPD (extrait de levure 10 g/l, peptone 20 g/l, glucose 20 g/l) contenant de la généticine à 0,2 g/l. Des souches transformées par le plasmide pYG72 ne contenant pas de cassette d'expression ont été sélectionnées pour servir de témoin dans les tests de production. Par ailleurs, afin de comparer l'efficacité du promoteur PGK de K.lactis selon l'invention par rapport à celui de S.cerevisiae, des souches transformées par le vecteur pYG19 ont également été sélectionnées. Le vecteur pYG19 est analogue au vecteur pYG621, sauf que le gène de l'albumine est sous le controle du promoteur PGK de S.cerevisiae (EP 361 991).

a) Analyse des ARNm:

Les cellules sont cultivées à 28°C en milieu sélectif YPD (extrait de levure 10 g/l, peptone 20 g/l, glucose 20 g/l) contenant de la généticine à 0,2 g/l. Les ARN totaux sont extraits (Sherman et al., Methods in Yeast Genetics, Cold Spring Harbor Laboratoty, 1986, 143) et séparés par électrophorèse sur gel d'agarose. Suivant la méthode de "Northern blot" (Maniatis et al., 1982 Molecular cloning, Cold Spring Harbor, Laboratory Press), les ARN sont hybridés à une sonde correspondant au gène de structure de l'albumine (fragment HindIII-HindIII de 1,9 kb) provenant du vecteur pYG18 (figure 7). L'autoradiographie montre clairement une bande de 2,3 kb spécifique de l'albumine (figure 13). Par ailleurs, il apparait clairement que le taux de transcription du gène de l'albumine est bien supérieur dans les souches contenant une région promotrice de l'invention (pYG621), que dans celles contenant le promoteur PGK intact de <u>S. cerevisiae</u> (pYG19).

b) Analyse des protéines :

Les cellules sont cultivées en erlenmeyers dans un milieu sélectif YPD (extrait de levure 10 g/l, peptone 20 g/l, glucose 20 g/l) contenant de la généticine à 0,2 g/l à 28°C sous agitation. Après 96 heures de culture, 30 µl de surnageant sont prélevés et mélangés à un volume équivalent de tampon Laemmli 2X (Laemmli, 1970, Nature 227, 680). Après chauffage à 96°C pendant 10 minutes, les protéines de l'échantil-Ion sont ensuite séparées sur gel de polyacrylamide SDS 8,5 %. La production d'albumine est ensuite révélée par coloration du gel au bleu de coomassie, puis est évaluée pour les différents vecteurs utilisés. La figure 14 montre que les 4 clones obtenus séparément par transformation de la souche MW98-8C par le vecteur pYG621 sécrètent beaucoup plus d'albumine que ceux obtenus par transformation avec le vecteur pYG19.

Il est clair que la région promotrice de l'invention permet une excellente production d'albumine par la levure, supérieure à celle obtenue avec le promoteur PGK de S.cerevisiae. Cette région, ou des formes réduites ou dérivées de celle-ci, constituent un outil industriel important pour les systèmes de production microbiologiques, et plus particulièrement eucaryotes

Revendications

- Séquence d'ADN comprenant tout ou partie de la séquence présentée à la figure 1 ou de son brin complémentaire, ou d'un dérivé de celles-ci, et possédant une activité de promoteur transcriptionnel.
- Séquence d'ADN selon la revendication 1 comprenant tout ou partie de la séquence présentée sur la figure 6(b).
- ADN recombinant comprenant une séquence d'ADN selon les revendications 1 ou 2.
- 4. ADN recombinant selon la revendication 3 caractérisé en ce qu'il contient en outre un ou plusieurs gènes de structure.
 - 5. ADN recombinant selon la revendication 4 caractérisé en ce qu'il contient également des signaux permettant la sécretion du produit d'expression du ou desdits gènes de structure.
 - 6. ADN recombinant selon l'une quelconque des re-

vendications 3 à 5 caractérisé en ce qu'il fait partie d'un plasmide d'expression, qui peut être à réplication autonome ou intégratif.

 Cellule recombinée contenant une séquence d'ADN ou un ADN recombinant selon l'une quelconque des revendications précédentes.

 Cellule recombinée selon la revendication 7 caractérisée en ce qu'il s'agit d'une levure.

ractérisée en ce qu'il s'agit d'une levure.

 Cellule recombinée selon la revendication 8 caractérisée en ce qu'il s'agit d'une levure du genre Kluyveromyces.

 Utilisation d'une séquence d'ADN selon l'une quelconque des revendications 1 à 6 pour l'expression de gènes recombinés.

11. Utilisation selon la revendication 10 pour l'expression simultanée, dans les 2 orientations opposées, de gènes recombinés.

 Utilisation selon l'une des revendications 10 ou 11 pour l'expression de gènes codant pour des protéines d'intérêt pharmaceutique ou agroalimentaire.

- 13. Procédé de préparation d'une protéine recombinante par expression de son gène dans un hôte cellulaire caractérisé en ce que l'expression dudit gène est sous le contrôle d'une séquence selon la revendication 1.
- Procédé selon la revendication 13 caractérisé en ce que la protéine est la sérum-albumine humaine.

5

15

10

20

25

30

35

40

45

50

EP 0 511 912 A1

TCTAGATTTA GCGGGTCATC GAAATTTAGT AGCGAGTCTA TTAGGGACCA GAGTTGCAAC CTGAGGTTTA ATGCGTCATC CTGTCGTTGC TTCAAGTTCC CCACTTGAAT CACTTGGACA 150 160 AACCGTTTCA TTGGTTTGAG GAAGGTGACG GATCTGGGTA GAAACTGGAC TACTGCATCT GTTGGTAGTC TTGATGCCAT GGTGATGAGC CATTGCCATT GGAAAAGAGT GAATTCAGAT TCCAAGATTT GGTCAATGAT TGATTTTGTA AGATTGAGAT CGTAATCCTG ATACTCTTTG AGCCATGTTT CCAACAGTTC TTCGGAATCT GCCGGTGTGG AAACGAGTAT TTCGGAGTAC 390 400 AATCTCGGTG GTTGCGTTAT CTGAGAGGAT GGTGTAGTGG TTTGATGTTG CTGTGTGAAA GATGATGCAG AGCTGATCAA CGATTCGAAC TGGGAGATCA CTTCGTTCAC TTCTTCCTGG 520 530 TTCCCGTTAC CTGTTTGCGT TTCCTCATAC ATTGGTACGC TATCCTCATC TTCAGATAAC GAAATATCAA ACTCATCGGA ATCGGACGCG TCGTTCAAAT CGCCCTCATC CTTGGTAATG 620 630 TTCTTGAACC GGTCGAGAAG GTTGAGAATC TCTGTCGGAA CACCACCCTG CGGCGTATAC CAGAACCAGA ATAAATTGTA GCACATCTTA ACTITCTCTA AGGAAACATC TGAACTCTGA TCAACGCATT CCGTAAGTAT ACTGTTTGCC TTGTCTCTGG TGAATTTATG AGGGTAAGAC TCTGAGATCA TAAGTAACTG TTGAGCATCG AAGTTGTTGT AGTTTGAAAT TAGGGATCTG GAAAGATGCG GTACCACTGC TTTGATGACA TTATCTGGCG GGTTCAACGG TACCAATTCC TGCAAGAATA GCGAATCCAA CGGTTTTAAC TCAGAGTAAT GGTTGATCAA CTCGATGAAA ACGTCCCAAT GGATGGATTG CATCAAGTGT TGATGTTCCA CCAAATTAAG ACAATATTTC 1030 . 1040 1050 GTAACGTTTT CGAGTGAAAC TGACACGGGC CTGCCCTCAG CACTCGTAGA CACGAGTAAC GTCTTGAGAC CTCTCGTACA GGGAAGCGAC ATATCGTTCA ATAGACTATG GAACAAAGTG TACACCGCAG CGATATCCTT GCATTTGCAA AACGATTGAA TAAGTGACGT CGATGCTAAA TCCTGGATAA GTACGCTGGT ATCGTGTAAG CCCATGAGAA CGACACGTTC CTCATCACTA

EP 0 511 912 A1

1270 GAAGCCGAAC	1280 TGTTGTCTTC	1290 AGTGGGGATT	1300 GGTTCGACAT	1310 TTTGCCAATT	1320 GCTGTCGATG	
1330 TACCCTTTCA	1340 AAGCCATGTA	1350 CCTTAAATCT	1360 TCATCCTTGG	1370 CAAGTAGATT	1380 CATCGGGTGT	
1390 GTTTGAAGTA	1400 AGAATATTTG	1410 CTTGTTTTTA	1420 TGGTATCAAA	1430 GGTATATGTT	1440 GTAGAAGACA	
1450 ATTTCCGGTA	1460 ATCCAATTGT	1470 CTGTCTGCTC	1480 AGTTTAGCAC	1490 ATGTATAGTA	1500 CGTTGCACAT	
1510 AGTCTACAAT	1520 ATTCAGCATT	1530 CAGCATTCAG	1540 TATACAGCAT	1550 ATGGCTAAAT	1560 GATCACAAAT	
1570 GTGATTGATG	1580 ATTTGACACG	1590 ACTAGAAAAG	1600 AGAACGAAAA	1610 AGGGAAATTC	1620 ATGTCACGTG	
1630 CGTTGGCACG	1640 TGACATGGAA	1650 TATCGAAGAA	1660 AGAAAAAAA	1670 AAACGATCTC	1680 GTCCTAGTGG	
1690	1700 TCTGGTCCCC	1710	1720	1730	1740	
1750 CAGAACACCC	1760 CACAGCAAAT	1770 GCACCACGCT	1780 ACGTAGATCA	1790 GGAAGCTTAA	1800 CTCTAGCGAC	
1810	1820 CCCCACAGAA	1830	1840	1850	1860	
1870 CTATACTCAT	1880 CTTGCTTCCC	1890 TTAAGCGTTC	1900 TCACGATTCG	1910 TTCGCTGCCC	1920 TTCTTCAAGA	
1930 GTCTTCTGAT	1940 TCTAATTCTC	1950 ATTCGAAATC	1960 CTCTACAGTT	 1970 AATGAATTGC	1980 TTGACATGAC	
1990	2000 TCATGGTTTT	2010	2020	2030	2040	
2050	2060 <u>TA</u> CGTCAAAA	2070	2080	2000	21.00	
2110	2120 TCAATCTATT	2130	2140	2150	21.60	
2170		21 90	2200	2210	2220	
2230	2240	2250				
AAAAAATCGA TCAAGAATTA ATAAAAAATG Met						

FIGURE 1 (suite)

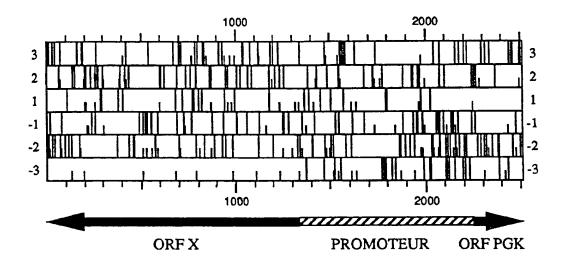


FIGURE 2

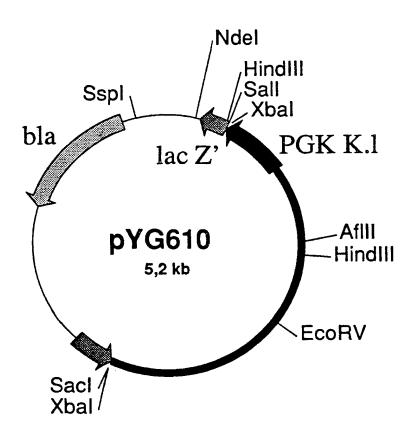


FIGURE 3

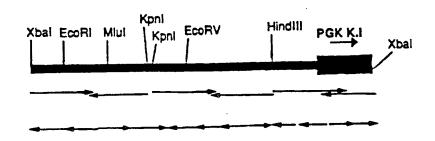


FIGURE 4

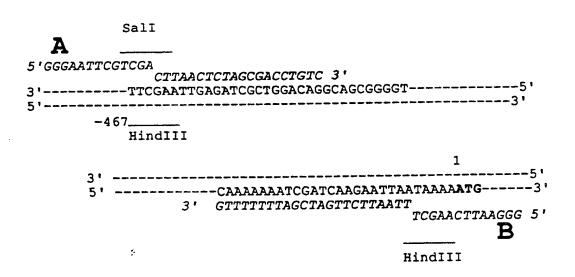


FIGURE 5

EP 0 511 912 A1

Oligodésoxynucléotide A
5'CATGTCGACTTTTTATTATTCTTGATCGAT3'
SalI
Oligodésoxynucléotide B
5'ATGAAGCTTAAATCTTCATCCTTGGC3'
HindIII
Oligodésoxynucléotide C
5'GGGTGAGGTTCTGTGGGGCGAGCGACAGGTCGCTAGAGTTAAGCATCCTGATC3'
(Position: 439 à 492)

Oligodésoxynucléotide D

5'GATCAGGATGCTTAACTCTAGCGACCTGTCGCTCGCCCCACAGAACCTCACCC3'

(Position: 439 à 492)

HindIII 10	- -	30 40)
aagcttTTAA ATCTTC					60
ttcgaaAATT TAGAAG					
TTTGCTTGTT TTTATG AAACGAACAA AAATAC	GTAT CAAAGGTAT CATA GTTTCCATA				120
TTGTCTGTCT GCTCAG AACAGACAGA CGAGTC					180
CATTCAGCAT TCAGTA GTAAGTCGTA AGTCAT	TACA GCATATGGC	T AAATGATCAC	AAATGTGATT	GATGATTTGA	240
CACGACTAGA AAAGAG GTGCTGATCT TTTCTC					300
GGAATATCGA AGAAAG CCTTATAGCT TCTTTC					360
CCCCCGGAG TCTTCC GGGGGGCCTC AGAAGG					420
AAATGCACCA CGCTAC TTTACGTGGT GCGATG					480
AGAACCTCAC CCGAGA TCTTGGAGTG GGCTCT					540
TCCCTTAAGC GTTCTC AGGGAATTCG CAAGAG					600
TCTCATTCGA AATCCT AGAGTAAGCT TTAGGA					660
TTTTGGCTTT TTGGCT AAAACCGAAA AACCGA	TTTG TCTTTTAAA AAAC AGAAAATTT	G CTATATCAAC C GATATAGTTG	TTTACATATA AAATGTATAT	AATATACGTC TTATATGCAG	720
AAAAGGGGAT TCATTA TTTTCCCCTA AGTAAT	ATTA GAAAATTCT TAAT CTTTTAAGA	C TTTTTCAATA G AAAAAGTTAT	GTTGCTATTC CAACGATAAG	ATTATCAATC TAATAGTTAG	780
TATTCAACTC AATTGG ATAAGTTGAG TTAACC		C TTTTTGTCAT G AAAAACAGTA			840
TAAATATATC TGTTGC ATTTATATAG ACAACG	TACA TTAAGAGTT ATGT AATTCTCAA	A CTTCAGAAAT T GAAGTCTTTA	AACAAAAAA TTGTTTTTT	TCGATCAAGA AGCTAGTTCT	900
ATTAATAAAA A <u>gtcga</u> TAATTATTTT Tcagct	g				917
10 SalI	20	30 4	0 50) 60)

FIGURE 6

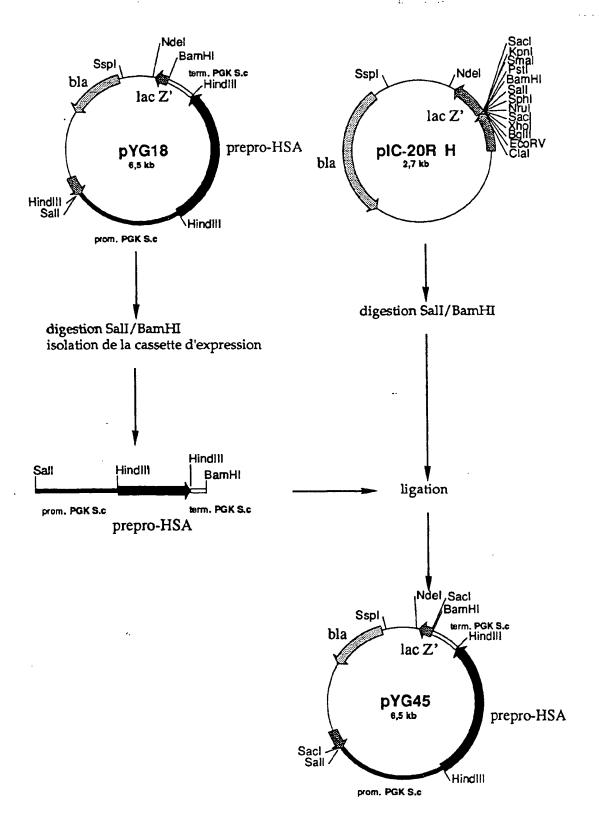


Figure 7

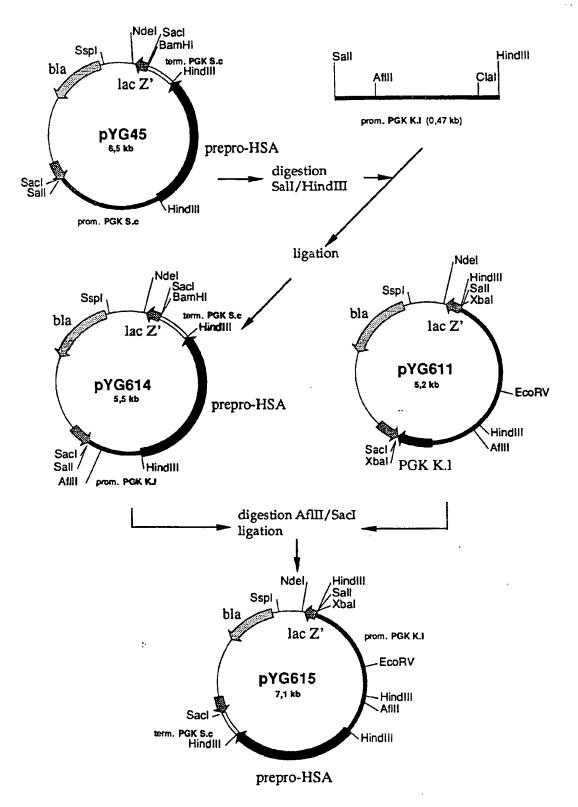


Figure 8

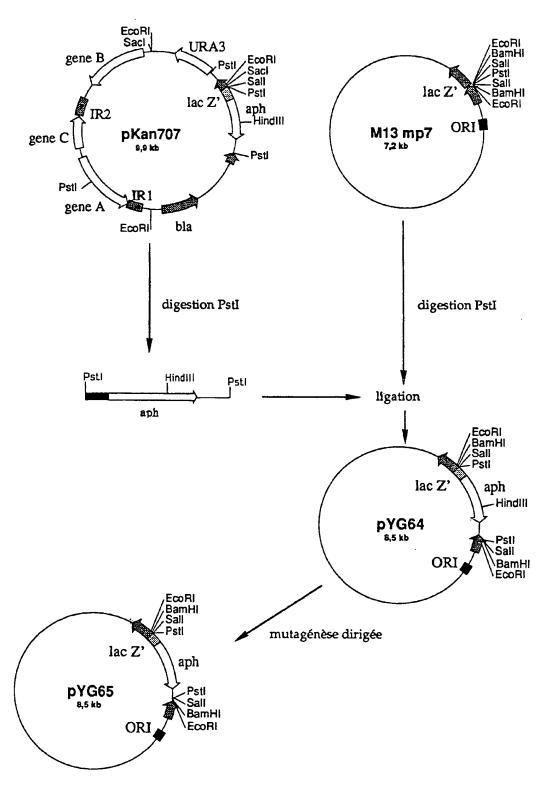


Figure 9

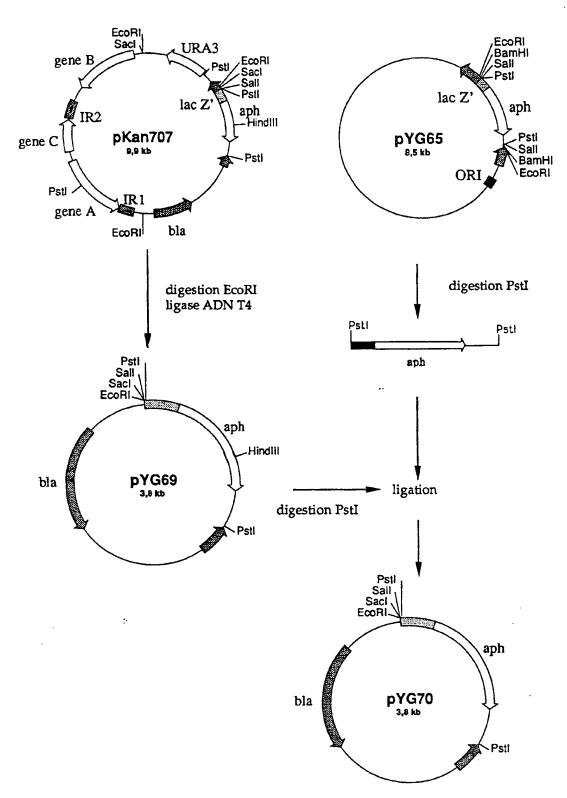


Figure 10

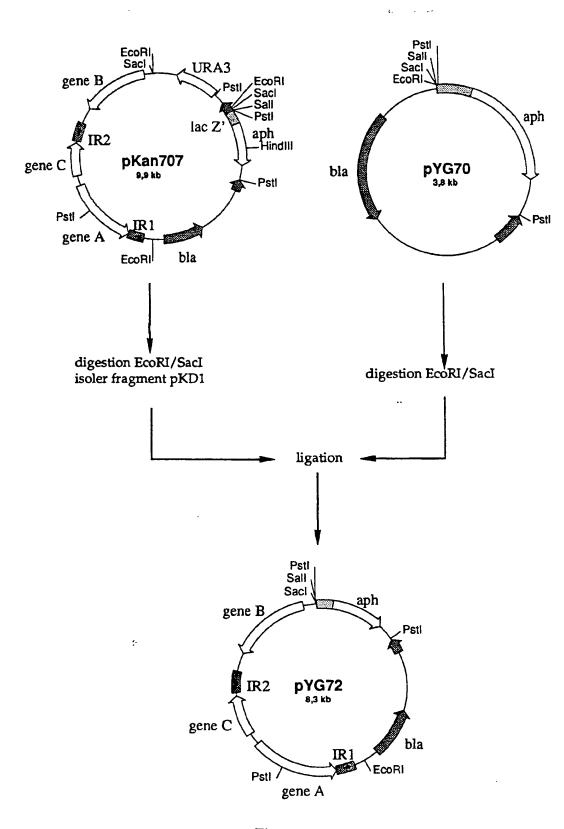


Figure 11

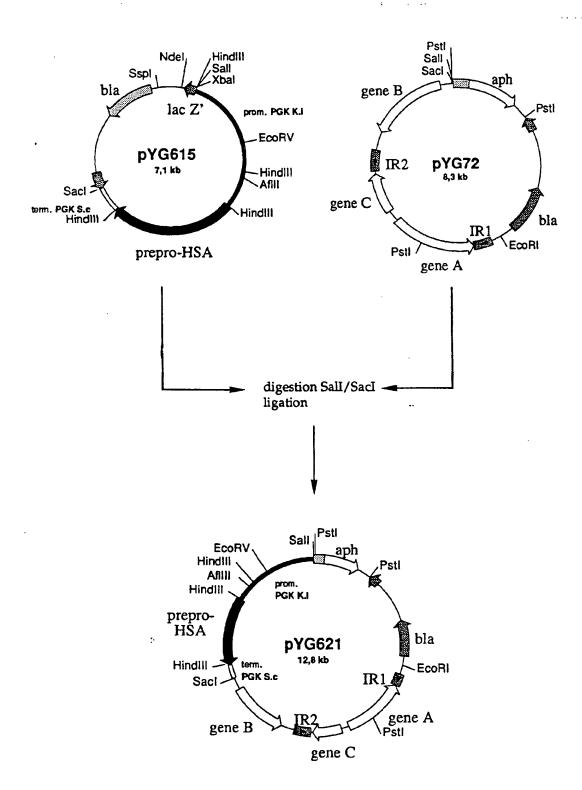


Figure 12

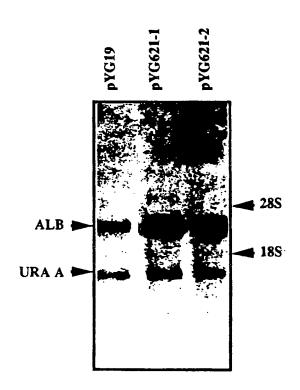


FIGURE 13

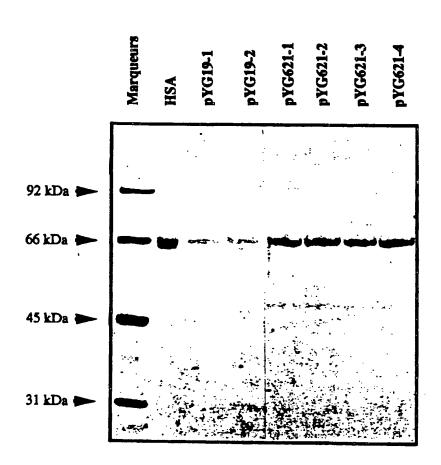


FIGURE 14



RAPPORT DE RECHERCHE EUROPEENNE Namero de la demande

EP 92 40 1206

tégorie	Citation du document avec ind		Revendication concernée	CLASSEMENT DE LA DEMANDE (lat. Cl.5)
	des parties pertin EP-A-0 361 991 (RHONE-POL 1990 * revendications 13-16 *		1-12	C12N15/81 C12N1/19 C12P21/U2
	₩		1	
	NUCLEIC ACIDS RESEARCH. vol. 18, no. 2, 25 Janvi virginia US page 365; A. FOURNIER, R. FLEER, P. 'The primary structure o kinase (PGK) gene from K * le document en entier	YEH AND JF, MAYAUX: f the 3-phosphoglycerate luyveromyces lactis!	1-12	·
Y	BIOTECHNOLOGY vol. 8, no. 2, Février 1 pages 135 - 139; J.A. VAN DEN BERG ET AL. host for heterologous ge expression and secretion * page 135, colonne 2, 1 colonne 2, 1 igne 39; fig	: 'Kluyveromyces as a me expression: of prochymosin' igne 21 - page 135,	1-12	DOMAINES TECHNIQUES RECHERCHES (Int. CL5)
A	JOURNAL OF BASIC MICROBI vol. 28, no. 4, 1988, 81 pages 211 - 220; X.J. CHEN ET AL.: 'A ge Kluyveromyces lactis an chromosomal gene requir production' * page 214, ligne 13 - 1	ERLIN, GERMANY ne-cloning system for d isolation of a ed for killer toxin		C12N C12P
La	présent rapport a été établi pour to	utes les revendications Duto d'achivement de la racherche		Dombales
X:1 Y:1 A:1		22 JUIN 1992	VA	N PUTTEN A.J.
A:	CATEGORIE DES DOCUMENTS oparticulièrement pertinent à lui seul particulièrement pertinent en combinaise surre document de la même catégorie unitère plan technologique divulgation non-écrite locument intercalaire	CITES T: théorie o E: éocumen éate és é n avec un D: cité éan L: cité pour	u principe à la base d t de bravet antérieur, épôt ou après cette d la dezrande d'autres raisons	e l'invention mais publié à la

This Page is Inserted by IFW Indexing and Scanning Operations and is not part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the i	tems checked:
☐ BLACK BORDERS	
☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES	
☐ FADED TEXT OR DRAWING	
BLURRED OR ILLEGIBLE TEXT OR DRAWING	¥
☐ SKEWED/SLANTED IMAGES	
☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS	
☐ GRAY SCALE DOCUMENTS	
☐ LINES OR MARKS ON ORIGINAL DOCUMENT	
☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR (QUALITY
	•

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.